

BC #5



PCT09

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/926,084

DATE: 01/24/2002
TIME: 12:50:12

Input Set : A:\ES.txt
Output Set: N:\CRF3\01242002\I926084.raw

ENTERED

3 <110> APPLICANT: SYODA, MAKOTO
4 SUGANO, YASUSHI
5 KUBOTA, HIDETOSHI
7 <120> TITLE OF INVENTION: NOVEL ENZYME WITH DECOLORIZING ACTIVITY AND METHOD FOR
DECOLORIZING DYES
8 BY USING THE SAME
10 <130> FILE REFERENCE: 213129US0PCT
12 <140> CURRENT APPLICATION NUMBER: US 09/926,084
13 <141> CURRENT FILING DATE: 2001-08-27
15 <150> PRIOR APPLICATION NUMBER: JP99/11-050562
16 <151> PRIOR FILING DATE: 1999-02-26
18 <160> NUMBER OF SEQ ID NOS: 16
20 <170> SOFTWARE: PatentIn version 3.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 6
24 <212> TYPE: PRT
25 <213> ORGANISM: Geotrichum candidum
27 <400> SEQUENCE: 1
29 Thr Tyr Val Pro Glu Arg
30 1 5
33 <210> SEQ ID NO: 2
34 <211> LENGTH: 8
35 <212> TYPE: PRT
36 <213> ORGANISM: Geotrichum candidum
38 <400> SEQUENCE: 2
40 Cys Pro Phe Gly Ala His Val Arg
41 1 5
44 <210> SEQ ID NO: 3
45 <211> LENGTH: 21
46 <212> TYPE: PRT
47 <213> ORGANISM: Geotrichum candidum
49 <400> SEQUENCE: 3
51 Ile Pro Tyr Gly Pro Glu Thr Ser Asp Ala Glu Leu Ala Ser Gly Val
52 1 5 10 15
55 Thr Ala Gln Asp Arg
56 20
59 <210> SEQ ID NO: 4
60 <211> LENGTH: 21
61 <212> TYPE: PRT
62 <213> ORGANISM: Geotrichum candidum
64 <400> SEQUENCE: 4
66 Ser Gly Ala Pro Ile Asp Leu Ala Pro Thr Ala Asp Asp Pro Ala Leu
67 1 5 10 15
70 Gly Ala Asp Pro Gln

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/926,084

DATE: 01/24/2002

TIME: 12:50:12

Input Set : A:\ES.txt

Output Set: N:\CRF3\01242002\I926084.raw

```

71          20
74 <210> SEQ ID NO: 5
75 <211> LENGTH: 6
76 <212> TYPE: PRT
77 <213> ORGANISM: Geotrichum candidum
79 <400> SEQUENCE: 5
81 Pro Tyr Gly Pro Glu Thr
82 1          5
85 <210> SEQ ID NO: 6
86 <211> LENGTH: 6
87 <212> TYPE: PRT
88 <213> ORGANISM: Geotrichum candidum
90 <400> SEQUENCE: 6
92 Pro Thr Ala Asp Asp Pro
93 1          5
96 <210> SEQ ID NO: 7
97 <211> LENGTH: 498
98 <212> TYPE: PRT
99 <213> ORGANISM: Geotrichum candidum
101 <400> SEQUENCE: 7
103 Met Asp Leu Ser Leu Phe Val Val Ser Val Ala Val Leu Val Gly Ser
104 1          5          10          15
107 Ser Ser His Val Asn Ala Ala Lys Leu Gly Ala Arg Gln Thr Arg Thr
108          20          25          30
111 Thr Pro Leu Leu Thr Asn Phe Pro Gly Gln Ala Pro Leu Pro Thr Leu
112          35          40          45
115 Thr Gln His Thr Thr Glu Ser Gly Ala Asn Asp Thr Ile Leu Pro Leu
116          50          55          60
119 Asn Asn Ile Gln Gly Asp Ile Leu Val Gly Met Lys Lys Gln Lys Glu
120 65          70          75          80
123 Arg Phe Val Phe Phe Gln Val Asn Asp Ala Thr Ser Phe Lys Thr Ala
124          85          90          95
127 Leu Lys Thr Tyr Val Pro Gln Arg Ile Thr Ser Ala Ala Ile Leu Ile
128          100          105          110
131 Ser Asp Pro Ser Gln Gln Pro Leu Ala Phe Val Asn Leu Gly Phe Ser
132          115          120          125
135 Asn Thr Gly Leu Gln Ala Leu Gly Ile Thr Asp Asp Leu Gly Asp Ala
136          130          135          140
139 Gln Phe Pro Asp Gly Gln Phe Ala Asp Ala Ala Asn Leu Gly Asp Asp
140 145          150          155          160
143 Leu Ser Gln Trp Val Ala Pro Phe Thr Gly Thr Thr Ile His Gly Val
144          165          170          175
147 Phe Leu Ile Gly Ser Asp Gln Asp Asp Phe Leu Asp Gln Phe Thr Asp
148          180          185          190
151 Asp Ile Ser Ser Thr Phe Gly Ser Ser Ile Thr Gln Val Gln Ala Leu
152          195          200          205
155 Ser Gly Ser Ala Arg Pro Phe Asp Gln Ala Gly His Glu His Phe Gly
156          210          215          220
159 Phe Leu Asp Gly Ile Ser Gln Pro Ser Val Thr Gly Trp Glu Thr Thr

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/926,084

DATE: 01/24/2002

TIME: 12:50:12

Input Set : A:\ES.txt

Output Set: N:\CRF3\01242002\I926084.raw

```

160 225          230          235          240
163 Val Phe Pro Gly Gln Ala Val Val Pro Pro Gly Ile Ile Leu Thr Gly
164          245          250          255
167 Arg Asp Gly Asp Thr Gly Thr Arg Pro Ser Trp Ala Leu Asp Gly Ser
168          260          265          270
171 Phe Met Ala Phe Arg His Phe Gln Gln Lys Val Pro Glu Phe Asn Ala
172          275          280          285
175 Tyr Thr Leu Ala Asn Ala Ile Pro Ala Asn Ser Ala Gly Asn Leu Thr
176          290          295          300
179 Gln Gln Glu Gly Ala Glu Phe Leu Gly Ala Arg Met Phe Gly Arg Trp
180 305          310          315          320
183 Lys Ser Gly Ala Pro Ile Asp Leu Ala Pro Thr Ala Asp Asp Pro Ala
184          325          330          335
187 Leu Gly Ala Asp Pro Gln Arg Asn Asn Phe Asp Tyr Ser Asp Thr
188          340          345          350
191 Leu Thr Asp Glu Thr Arg Cys Pro Phe Gly Ala His Val Arg Lys Thr
192          355          360          365
195 Asn Pro Arg Gln Asp Leu Gly Gly Pro Val Asp Thr Phe His Ala Met
196          370          375          380
199 Arg Ser Ser Ile Pro Tyr Gly Pro Glu Thr Ser Asp Ala Glu Leu Ala
200 385          390          395          400
203 Ser Gly Val Thr Ala Gln Asp Arg Gly Leu Leu Phe Val Glu Tyr Gln
204          405          410          415
207 Ser Ile Ile Gly Asn Gly Phe Arg Phe Gln Gln Ile Asn Trp Ala Asn
208          420          425          430
211 Asn Ala Asn Phe Pro Phe Ser Lys Pro Ile Thr Pro Gly Ile Glu Pro
212          435          440          445
215 Ile Ile Gly Gln Thr Thr Pro Arg Thr Val Gly Gly Leu Asp Pro Leu
216          450          455          460
219 Asn Gln Asn Glu Thr Phe Thr Val Pro Leu Phe Val Ile Pro Lys Gly
220 465          470          475          480
223 Gly Glu Tyr Phe Phe Leu Pro Ser Ile Ser Ala Leu Thr Ala Thr Ile
224          485          490          495
227 Ala Ala
231 <210> SEQ ID NO: 8
232 <211> LENGTH: 1494
233 <212> TYPE: DNA
234 <213> ORGANISM: Geotrichum candidum
236 <400> SEQUENCE: 8
237 atgcgcttgt cgctgtttgt cgtgtcggtt gccgtactcg tcgggtcgag ctgcgatgtc      60
239 aatgctgcta aactcggcgc gagacagacg cgtacgacac ccctcctcac taattttccg      120
241 ggacaagccc cgctgccgac tctaacgcag catacgactg agagcggggc caacgataca      180
243 attctgcccc tgaacaacat acaaggcgac attttggttg gcatgaagaa acagaaggaa      240
245 cgcttcgtct ttttccaagt caatgacgca acctcgttca agacggcggt gaagacctac      300
247 gtgcctgagc gcatcacgtc ggcggcgatt ttgatttcag atccttctca gcagccgttg      360
249 gctttcgtca acctcgggtt ttcgaacaca ggctccagg cgcttggaat taccgacgat      420
251 ctgggtgatg cacaattccc agatggtcag ttcgcagacg ccgcaaacct cggggacgac      480
253 ctcagccaat ggggtggcgc ttttactggt accaccatcc atggtgtctt tctgattggt      540
255 agcgaccagg acgacttctt ggatcagttc acggatgata tctcttcgac ctttggttcc      600

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/926,084

DATE: 01/24/2002

TIME: 12:50:12

Input Set : A:\ES.txt

Output Set: N:\CRF3\01242002\I926084.raw

```

257 tccatcactc aggtgcaggc gctcagtggg tctgcgcgtc caggagatca ggctgggtcat 660
259 gaacacttcg ggttcctcga cggcatctcg cagccctcag tcacaggctg ggagacgacc 720
261 gtcttccctg gacaggcggg cgtccacact ggaattatcc tcaactggac cgatggggac 780
263 acgggcaccc gacgctcgtg ggctctagat gggagtttca tggcattccg gcacttccag 840
265 cagaaggtcc ccgaattcaa cgcgtacacg ctgcgccaacg cgatacccg cgaacagcgcg 900
267 ggaaacctca cccagcagga aggtgcagag ttctctcgcg cgcgcagtgtt cggccggttg 960
269 aagagcgggc cgccgattga cctcgcgcgc acggcgggac acccagcgct cggcgccgac 1020
271 ccgcagagga acaacaattt cgattactca gacacgctga cggacgagac gcgctgcccc 1080
273 ttcggtgcac acgtgaggaa gacgaaccct cgacaggacc tgggtggacc ggtcgacacc 1140
275 ttccacgcta tgcggtccag tatcccgtac ggcccagaaa cgtctgatgc agaacttgcg 1200
277 tcgggcggtg ctgcgcaaga ccgcggtctt cttttcgtcg agtaccagtc cattattggt 1260
279 aatgggttca ggttccagca gattaactgg gcgaacaatg cgaacttccc tttctccaaa 1320
281 ccgatcacgc ctggaattga gcctatcatc ggccagacga ctccacgcac tgcgcgggg 1380
283 ctgcaccccc tcaaccagaa tgagacgttc acagtaccgc tgtttgatgat cccgaagggc 1440
285 ggggaatact ttttcttgcc ctctatctct gcgctcactg cgactatcgc tgct 1494

```

288 <210> SEQ ID NO: 9

289 <211> LENGTH: 26

290 <212> TYPE: PRT

291 <213> ORGANISM: Geotrichum candidum

293 <220> FEATURE:

294 <221> NAME/KEY: NON_CONS

295 <222> LOCATION: (13)..(14)

296 <223> OTHER INFORMATION:

299 <400> SEQUENCE: 9

301 Gln Ala Pro Leu Pro Thr Leu Thr Gln His Thr Thr Glu Val Ala Pro

302 1 5 10 15

305 Phe Thr Gly Thr Thr Ile His Gly Val Phe

306 20 25

309 <210> SEQ ID NO: 10

310 <211> LENGTH: 26

311 <212> TYPE: PRT

312 <213> ORGANISM: Geotrichum candidum

314 <220> FEATURE:

315 <221> NAME/KEY: NON_CONS

316 <222> LOCATION: (13)..(14)

317 <223> OTHER INFORMATION:

320 <400> SEQUENCE: 10

322 Gly Pro Val Leu Val Arg Leu Ala Trp His Thr Ser Gly Arg Glu Val

323 1 5 10 15

326 Val Ala Leu Met Gly Ala His Ala Leu Gly

327 20 25

330 <210> SEQ ID NO: 11

331 <211> LENGTH: 26

332 <212> TYPE: PRT

333 <213> ORGANISM: Geotrichum candidum

335 <220> FEATURE:

336 <221> NAME/KEY: NON_CONS

337 <222> LOCATION: (13)..(14)

338 <223> OTHER INFORMATION:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/926,084

DATE: 01/24/2002

TIME: 12:50:12

Input Set : A:\ES.txt

Output Set: N:\CRF3\01242002\I926084.raw

341 <400> SEQUENCE: 11

343 Ala Gly Leu Phe Ile Arg Met Ala Trp His Gly Ala Gly Glu Thr Val

344 1 5 10 15

347 Ala Leu Ile Ala Gly Gly His Thr Leu Gly

348 20 25

351 <210> SEQ ID NO: 12

352 <211> LENGTH: 26

353 <212> TYPE: PRT

354 <213> ORGANISM: Geotrichum candidum

356 <220> FEATURE:

357 <221> NAME/KEY: NON_CONS

358 <222> LOCATION: (13)..(14)

359 <223> OTHER INFORMATION:

362 <400> SEQUENCE: 12

364 Val Arg Lys Ile Leu Arg Ile Val Phe His Asp Ala Ile Asp Glu Val

365 1 5 10 15

368 Val Asp Leu Leu Ala Ala His Ser Leu Ala

369 20 25

372 <210> SEQ ID NO: 13

373 <211> LENGTH: 26

374 <212> TYPE: PRT

375 <213> ORGANISM: Geotrichum candidum

377 <220> FEATURE:

378 <221> NAME/KEY: NON_CONS

379 <222> LOCATION: (13)..(14)

380 <223> OTHER INFORMATION:

383 <400> SEQUENCE: 13

385 Ala His Glu Val Ile Arg Leu Thr Phe His Asp Ala Ile Phe Glu Val

386 1 5 10 15

389 Val Ser Leu Leu Ala Ser His Thr Val Ala

390 20 25

393 <210> SEQ ID NO: 14

394 <211> LENGTH: 26

395 <212> TYPE: PRT

396 <213> ORGANISM: Geotrichum candidum

398 <220> FEATURE:

399 <221> NAME/KEY: NON_CONS

400 <222> LOCATION: (13)..(14)

401 <223> OTHER INFORMATION:

404 <400> SEQUENCE: 14

406 Ala His Glu Ser Ile Arg Leu Val Phe His Asp Ser Ile Leu Glu Leu

407 1 5 10 15

410 Val Trp Met Leu Ser Ala His Ser Val Ala

411 20 25

414 <210> SEQ ID NO: 15

415 <211> LENGTH: 26

416 <212> TYPE: PRT

417 <213> ORGANISM: Geotrichum candidum

419 <220> FEATURE:

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/926,084

DATE: 01/24/2002

TIME: 12:50:13

Input Set : A:\ES.txt

Output Set: N:\CRF3\01242002\I926084.raw